A story... of one convex relaxation... and of the related revelation...

Leonid Gurvits

Los Alamos National Laboratory , Nuevo Mexico, USA.
e-mail: gurvits@lanl.gov

The Mixed volume and the Mixed Discriminant, 1998, A. Barvinok's paper in "Lectures on Mathematical Programming: ISMP-97"

$\mathbf{K} = (K_1, ..., K_n)$ is a n-tuple of convex compact subsets (i.e. **convex bodies**) in the Euclidean space R^n;

$$V_{\mathbf{K}}(\lambda_1, ..., \lambda_n) =: Vol(\lambda_1 K_1 + \cdots + \lambda_n K_n), \lambda_i \geq 0.$$

Herman Minkowski proved that $V_{\mathbf{K}}$ is a homogeneous polynomial with non-negative coefficients. The mixed volume:

$$V(K_1, ..., K_n) =: \frac{\partial^n}{\partial \lambda_1 ... \partial \lambda_n} V_{\mathbf{K}}(0, ..., 0).$$

i.e. the mixed volume $V(K_1, ..., K_n)$ is the coefficient of the monomial $\Pi_{1 \leq i \leq n} \lambda_i$ in the Minkowski polynomial $V_{\mathbf{K}}$.

Let $\mathbf{A} = (A_1, ..., A_n)$ be an n-tuple of $n \times n$ complex matrices;

the corresponding determinantal polynomial is defined as

$$Det_{\mathbf{A}}(\lambda_1, ..., \lambda_n) = \det(\sum_{1 \leq i \leq n} \lambda_i A_i).$$

The mixed discriminant is defined as

$$D(A_1, ..., A_n) = \frac{\partial^n}{\partial \lambda_1 ... \partial \lambda_n} Det_{\mathbf{A}}(0, ..., 0).$$

i.e. the mixed discriminant $D(A_1, ..., A_n)$ is the coefficient of the monomial $\Pi_{1 \leq i \leq n} \lambda_i$ in the determinantal polynomial $Det_{\mathbf{A}}$.

Examples.

1.

$$K_i = \{(t_1, ..., t_n) : 0 \le t_j \le A(i,j), 1 \le j \le n\},$$

the mixed volume of coordinate boxes K_i:

$$V(K_1...K_n) = Per(A) = \sum_{\sigma \in S_n} \prod_{1 \le i \le n} A(i, \sigma(i)).$$

If each coordinate box K_i is a rectangle(parallelogram) then computing the mixed volume $V(B_1, ..., B_n)$ is "easy".

2. $K_i = \{ae_i + bY_i : 0 \le a, b \le 1\}$ is a *parallelogram*, $A =: [Y_1, ..., Y_n]$. Then the mixed volume

$$V(K_1, ..., K_n) = MV(A) =: \sum_{S \subset \{1,...,n\}} |\det(A_{S,S})|.$$

If $Q_i = e_i e_i^T + Y_i Y_i^T$ then the mixed discriminant

$$D(Q_1, ..., Q_n) = MD(A) =: \sum_{S \subset \{1,...,n\}} (\det(A_{S,S}))^2.$$

3. Both $MD(A)$ and $MV(A)$ are $\#P - Complete$ even if the matrix A is **unimodular**.

From the mixed volume of Ellipsoids to the Mixed Discriminant

The **convex bodies** K_i are well-presented: Given a weak membership oracle for K_i and a rational $n \times n$ matrix B_i, a rational vector $Y_i \in R^n$ such that

$$Y_i + B_i(Ball_n(1)) \subset K_i \subset Y_i + n\sqrt{n+1}B_i(Ball_n(1)) \tag{1}$$

Let $\mathcal{E}_B$ be the ellipsoid $B(Ball_n(1))$ in R^n. Then

$$V(\mathcal{E}_{B_1}, ..., \mathcal{E}_{B_n}) \leq V(K_1, ..., K_n) \leq (n\sqrt{n+1})^n V(\mathcal{E}_{B_1}, ..., \mathcal{E}_{B_n}).$$

[Barvinok, 1997]: Define $v_n =: Vol_n(Ball_n(1))$. Then the following inequalities hold:

$$3^{-\frac{n+1}{2}} v_n D^{\frac{1}{2}}(A_1(A_1)^T, ..., A_n(A_n)^T) \leq V(\mathcal{E}_{A_1} ... \mathcal{E}_{A_n}) \leq v_n D^{\frac{1}{2}}(..) \tag{2}$$

Suppose that we have an effectively computable estimate F such that

$$\gamma(n) \leq \frac{D(A_1(A_1)^T, ..., A_n(A_n)^T))}{F} \leq 1.$$

Then

$$\sqrt{\gamma(n)} 3^{-\frac{n+1}{2}} \leq \frac{V(K_1, ..., K_n)}{\sqrt{F} v(n)} \leq n^{1.5n}$$

Which gives the approximation factor
$n^{1.5n} 3^{\frac{n+1}{2}} (\sqrt{\gamma(n)})^{-1} \geq n^{O(n)}$.
Barvinok [1997] gave the poly-time randomized algorithm with $\gamma(n) = c^n, c < 1$.

A deterministic algorithm for the Mixed Discriminant, Geometric Programming, Quantum Entanglement: 1998-2005

Let $p \in Hom_+(n,n)$ be a homogeneous polynomial with nonnegative coefficients. Define the following quantity, called **Capacity**:

$$Cap(p) =: \inf_{x_i > 0} \frac{p(x_1, \ldots, x_n)}{\Pi_{1 \leq i \leq n} x_i}.$$

Clearly $\frac{\partial^n}{\partial x_1 \cdots \partial x_n} p(0,0,...,0) \leq Cap(p)$.

Now,

$$\log(Cap(p)) = \inf_{y_1 + \ldots + y_n = 0} \log(p(e^{y_1}, ..., e^{y_n}))$$

and the functional $\log(p(e^{y_1}, ..., e^{y_n}))$ is convex. Therefore $\log(Cap(p))$ might be, with some extra care and luck, effectively additively approximated using convex programming tools and an oracle, deterministic or random, evaluating the polynomial p.

But we need a lower bound:

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} p(0, 0, ..., 0) \geq \gamma(n) Cap(p), \gamma(n) > 0.$$

In the case of the mixed discriminant the corresonding polynomial

$p(x_1, ..., x_n) = \det(\Sigma_{1 \leq i \leq n} x_i Q_i)$, where the matrices $Q_i \succeq 0$ are PSD. Easy to evaluate deterministically!

Boils down to the following result:

Theorem 0.1: *Let n-tuple $\mathbf{A} = (A_1, \ldots, A_n)$ of hermitian $n \times n$ PSD matrices be doubly-stochastic:*

$$tr(A_i) = 1, 1 \leq i \leq n; \sum_{1 \leq i \leq n} A_i = I.$$

Then the mixed discriminant

$$D(\mathbf{A}) =: \frac{\partial^n}{\partial x_1, \ldots, \partial x_n} Det_{\mathbf{A}}(0, \ldots, 0) \geq \frac{n!}{n^n} \qquad (3)$$

The equality in (3) is attained iff $A_i = \frac{1}{n} I, 1 \leq i \leq n$.

Solution of R. Bapat's conjecture (1989), stated for real symmetric PSD matrices, generalization of Van der Waerden conjecture for the permanent; proved by L.G. (1999), final publication (2006))

The reason for the result: optimality condition for $\min_{y_1 + \ldots + y_n = 0} \log(Det_{\mathbf{Q}}(e^{y_1}, ..., e^{y_n}))$ states that the tuple $(P(e^{y_1} Q_1)P, ..., P(e^{y_n} Q_n)P), P = (\Sigma_{1 \leq i \leq n} e^{y_i} Q_i)^{-\frac{1}{2}}$

is doubly-stochastic. This observation and Theorem(0.1) imply that

$$\frac{n!}{n^n} \leq \frac{D(Q_1, ..., Q_n)}{Cap(Det_{\mathbf{Q}})} \leq 1.$$

Can put $\gamma(n) = \frac{n!}{n^n} \approx e^{-n}$.

My proof is a very non-trivial adaptation of Egorychev's proof of Van Der Waerden conjecture for the permanent, which I learned from Knut's 1981 *Monthly* exposition.

Did not actually need doubly-stochasticity, it served as a tool; non-convex optimization with semi-definite constraints.

The proof is very matrix-oriented, crucially uses the group action:

$$D(XA_1X^*, ..., XA_nX^*) = \det(XX^*)D(A_1, ..., A_n).$$

Got a deterministic poly-time(not strongly polynomial) algorithm to approximate the **mixed discriminant** with the factor e^n and the mixed volume with the factor $n^{O(n)}$. Can we get a factor c^n deterministically for the **mixed volume**? NO!

In the oracle setting, even for the single volume the factor is greater than $\left(\Omega\left(\frac{n}{\log n}\right)\right)^{\frac{n}{2}}$ (**Barany-Furedi bound**).

Can we get factor c^n using a randomized poly-time algorithm?

Can we get a better factor for the **mixed discriminant** if the ranks $Rank(Q_i)$ are small?

Is there a simpler proof?

A revelation, 2003-2004-2005-...

Definition 0.2 A homogeneous polynomial $p \in Hom_C(m, n)$ is **H-Stable** if

$$|p(z_1, ..., z_m)| > 0; Re(z_i) > 0, 1 \leq i \leq m.$$

∎

Example 0.3 Consider a bivariate homogeneous polynomial $p(z_1, z_2) = (z_2)^n P(\frac{z_1}{z_2})$, where P is some univariate polynomial. Then p is **H-Stable** iff the roots of P are non-positive real numbers. This assertion is just a rephrasing of the next set equality:

$$\mathbf{C} - \{\frac{z_1}{z_2} : Re(z_1), Re(z_2) > 0\} = \{x \in R : x \leq 0\}.$$

∎

This simple bivariate observation gives the connection between **H-Stability** and *Hyperbolicity*:

Fact 0.4 A homogeneous polynomial $p \in Hom_C(m,n)$ is **H-Stable** iff it is e-hyperbolic, $e = (1, ..., 1)$, i.e. the roots of $p(x_1 - t, ..., x_m - t) = 0$ are real for all real vectors $X \in R^m$, and its hyperbolic cone contains the positive orthant R^m_{++}, i.e. the roots of $p(X - te) = 0$ are positive real numbers for all positive real vectors $X \in R^m_{++}$.
Moreover $\frac{p}{p(X)} \in Hom_+(m,n)$ for all $X \in R^m_{++}$ and

$|p(z_1, ..., z_m)| \geq |p(Re(z_1), ..., Re(z_m))| : Re(z_i) \geq 0,$
$1 \leq i \leq m.$

Note that a determinantal polynomial $Det_{\mathbf{Q}}$ is **H-Stable** for non-trivial PSD tuples:
$Q_i \succeq 0, \Sigma_{1 \le i \le n} Q_i \succ 0$.

A homogeneous polynomial $q \in Hom_{+}(n,n)$ is called doubly-stochastic if

$$\frac{\partial}{\partial x_i} q(1,1,\ldots,1) = 1, 1 \le i \le n.$$

Alternative definition:

$$q(x_1, ..., x_n) \ge \prod_{1 \le i \le n} x_i, x_i > 0; q(e) = 1. \qquad (4)$$

A determinantal polynomial $Det_{\mathbf{Q}}$ is **H-Stable** and doubly-stochastic for doubly-stochastic tuples $(Q_1, ..., Q_n)$ is doubly-stochastic !?!?... A possible generalization of Van Der Waerden and Bapat's conjectures, but how to prove it?

All previous proofs heavily relied on the matrix structure.

The **Capacity**, which appeared as an algorithmic tool, happened to be the “saviour”!

Theorem 0.5: *Let* $p \in Hom_+(n,n)$ *be* **H-Stable** *polynomial and*

$$G(i) = \left(\frac{i-1}{i}\right)^{i-1}, i > 1; G(1) = 1.$$

Then the following inequality holds

$$1 \geq \frac{\frac{\partial^n}{\partial x_1 \ldots \partial x_n} p(0,\ldots,0)}{Cap(p)} \geq \prod_{2 \leq i \leq n} G(\min(i, deg_p(i))). \tag{5}$$

Actually, $G(i) = \frac{wdv(i)}{wdv(i-1)}$, where $vdw(i) = \frac{i!}{i^i}$ and this function G is strictly decreasing on $[0, \infty)$.

Thus $\Pi_{2 \leq i \leq n} G(\min(i, deg_p(i))) \geq G(2) \cdots G(n) = \frac{n!}{n^n}$

Corollary 0.6: *Assume WLOG that* $Cap(p) > 0$. *Then*

$$\frac{\frac{\partial^n}{\partial x_1 \ldots \partial x_n} p(0,\ldots,0)}{Cap(p)} \geq \frac{n!}{n^n} \tag{6}$$

Equality in (6) is attained iff $p(x_1, ..., x_n) = (a_1 x_1 + ... + a_n x_n)^n; a_i > 0, 1 \leq i \leq n$.

Proof: Step 1.

Lemma 0.7: *Consider an univariate polynomial* $R(t) = \Sigma_{0 \leq i \leq k} a_i t^i; a_i \geq 0, a_k > 0$. *If the roots of* R *are real(nec. non-positive) then*

$$R'(0) \geq G(k) \inf_{t>0} \frac{R(t)}{t}, G(k) = \left(\frac{k-1}{k}\right)^{k-1} \qquad (7)$$

Proof: The case $R(0) = 0$ is trivial: $G(k) \geq 1$ and $R'(0) = \inf_{t>0} \frac{R(t)}{t}$.

Otherwise, $R(t) = R(0) \Pi_{1 \leq i \leq k}(1 + b_i t)$ where $b_i > 0, 1 \leq i \leq k$.

Assume WLOG that $R(0) = 1$. We get, using AM/GM inequality, that

$$R(t) \leq Pow(t) =: (1 + \frac{R'(0)t}{k})^k$$

Easy to compute that $\inf_{t>0} \frac{Pow(t)}{t} = R'(0)(G(k))^{-1}$.

Which leads to

$$R'(0)(G(k))^{-1} = \inf_{t>0} \frac{Pow(t)}{t} \geq inf_{t>0} \frac{R(t)}{t}.$$

Step 2.

Fix positive numbers $x_1, ..., x_{n-1}$ and consider the univariate polynomial $R(t) = p(x_1, ..., x_{n-1}, t)$. Note that the bivariate homogeneous polynomial $Q(s,t) = p(sx_1, ..., sx_{n-1}, t) = s^n R(t/s)$ is **H-Stable**. Therefore, the roots of R are non-positive real numbers. The degree $deg(R) = deg_p(n) =: k$ and

$$R(t) = p(x_1, ..., x_{n-1}, t) \geq Cap(p)(\prod_{1 \leq i \leq n-1} x_i)t, t \geq 0.$$

Lemma(0.7) gives the inequality

$$R'(0) \geq (Cap(p) \prod_{1 \leq i \leq n-1} x_i)G(deg_p(n)).$$

Note that $R'(0) = \frac{\partial}{\partial x_n} p(x_1, \ldots, x_{n-1}, 0) =: q_{n-1}(x_1, \ldots, x_{n-1})$. We finally get the main inequality:

$$Cap(q_{n-1}) \geq G(deg_p(n))Cap(p)$$

Step 3. Define the following polynomials $q_i \in Hom_+(i,i), 1 \leq i \leq n-1$:

$$q_i(x_1,\dots,x_i) = \frac{\partial^{n-i}}{\partial x_{i+1}\dots\partial x_n} p(x_1,\dots,x_i,0,\dots,0).$$

Note that $q_1(x_1) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} p(0) x_1, Cap(q_1) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} p(0)$ and $deg_{q_i}(i) \leq \min(i, deg_p(i))$.

Using Gauss-Lukas Theorem, we get that q_i is either zero or **H-Stable**. **Step 2** gives that

$$Cap(q_{i-1}) \geq G(deg_{q_i}(i)) Cap(q_i).$$

Since $deg_{q_i}(i) \leq \min(i, deg_p(i))$ abd G is decreasing, we get that

$$Cap(q_{i-1}) \geq G(\min(i, deg_p(i)) Cap(q_i). \tag{8}$$

Finally we just multiply inequalities (8):

$$\frac{\partial^n}{\partial x_1 \dots \partial x_n} p(0) = Cap(q_1) \geq Cap(p) \prod_{2 \leq i \leq n} G(\min(i, deg_p(i))).$$

∎

Specializing to the permanent and the mixed discriminant: the polynomial for permanent $Per(A)$ is

$$Prod_A(x_1, ..., x_n) = \prod_{1 \leq i \leq n} \sum_{1 \leq j \leq n} A(i,j)x_j.$$

I.e. the mixed derivative of $Prod_A$ is equal to $Per(A)$. If A is non-negative and $Prod_A \neq 0$ then $Prod_A$ is **H-Stable**.

$deg_{Prod_A}(j) = |col(j)| =$ number of non-zero entries in jth column.

If A is doubly-stochastic then $Cap(Prod_A) = 1$.

Theorem 0.8: *If A is a doubly-stochastic $n \times n$ matrix then*

$$Per(A) \geq \prod_{2 \leq j \leq n} G(\min(|col(j)|, j)) \geq \prod_{2 \leq i \leq n} G(j) = \frac{n!}{n^n}.$$

If $|col(j)| \leq k < n$ for $k + 1 \leq j \geq n$ then

$$Per(A) \geq \left(\frac{k-1}{k}\right)^{(k-1)(n-k)} \frac{k!}{k^k} > \left(\left(\frac{k-1}{k}\right)^{k-1}\right)^n \quad (9)$$

A. Schrijever (1998): $A = \{\frac{d(i,j)}{n} : 1 \leq i, j \leq n\}$, All rows and columns of the **integer** matrix D sum to $k \leq n$ (i.e. k-regular bipartite graph with multiple edges). Then

$$Per(A) \geq \left(\frac{k-1}{k}\right)^{(k-1)n}. \tag{10}$$

The inequality (9) gives a stronger version of the very discrete Schrijvers's inequality (10). Moreover, our inequality works in much more general real valued case. Amazingly, the exponent $\left(\frac{k-1}{k}\right)^{k-1}$ is optimal. This optimality follows from a forgotten H. Wilf's 1966 paper. Was rediscovered by Schrijver and Valiant in 1981.

In the case of the mixed discriminant of doubly-stochastic tuples:

$$D(A_1, ..., A_n) \geq \prod_{2 \leq j \leq n} G(\min(Rank(A_j), j)).$$

This leads to the deterministic poly-time algorithms to approximate as $\Sigma_{S \subset \{1,...,n\}} |\det(A_{S,S})|$ as well

$\Sigma_{S\subset\{1,\ldots,n\}} |\det(A_{S,S})|^2$ with the factor $\frac{2^n}{n^m}$.

Back to the mixed volume: the Minkowski polynomial

$$Vol_n(\lambda_1 K_1 + \cdots + \lambda_n K_n)) = V_{\mathbf{K}}(\lambda_1, ..., \lambda_n)$$

is not necessary **H-Stable** if $n \geq 3$. But essentially the same proof works!

Theorem 0.9: *Let* $\mathbf{K} = (K_1, ..., K_n)$ *be a tuple of conbex bodies in* R^n*. Then the mixed volume*

$$Cap(V_{\mathbf{K}}) \geq V(K_1, ..., K_n) \geq \frac{n!}{n^n} Cap(V_{\mathbf{K}}). \qquad (11)$$

The inequalities (11) lead to a randomized poly-time algorithm to approximate the mixed volume with the factor e^n.

Why it works: the polynomials

$$q_i(x_1, \ldots, x_i) = \frac{\partial^{n-i}}{\partial x_{i+1} \ldots \partial x_n} p(x_1, \ldots, x_i, 0, \ldots, 0)$$

are not **H-Stable**, but $(q_i)^{\frac{1}{i}}$ are log-concave on R^i_+. A lot of hyperbolic (i.e. **H-Stable**) stuff can generalized to such **Strongly Log-Concave** polynomials and even entire functions.

A few open problems

1. Let $p \in Hom_+(n,n)$ be **H-Stable** and doubly-stochastic, $Z = (z_1, ..., z_n) \in C^n$.

   Is the vector $\Gamma = (\gamma_1, ..., \gamma_n)$ consisting of all the roots of the equation $p(z_1 - t, ..., z_n - t) = 0$ majorized by the vector Z, i.e. $\Gamma = AZ$ for some DS matrix A ?

2. Let us consider two **H-Stable** polynomials $p, q \in Hom_+(m,n)$:

   $p(x_1, ..., x_m) = \Sigma_{r_1+...+r_m=n}\, a_{r_1,...,r_m} \Pi_{1 \leq i \leq m}\, x_i^{r_i}$,

   $q(x_1, ..., x_m) = \Sigma_{r_1+...+r_m=n}\, b_{r_1,...,r_m} \Pi_{1 \leq i \leq m}\, x_i^{r_i}$,

   and a nonnegative vector $(l_1, ..., l_m)$ with $\Sigma_{1 \leq i \leq m}\, l_i = n$.

   Let us assume that

   $\inf_{x_i > 0, 1 \leq i \leq m} \frac{p(x_1,...,x_m)}{\Pi_{1 \leq i \leq m}\, x_i^{l_i}} =: A > 0$,

   $\inf_{x_i > 0, 1 \leq i \leq m} \frac{q(x_1,...,x_m)}{\Pi_{1 \leq i \leq m}\, x_i^{l_i}} =: B > 0$.

Then the following inequality holds:

$$< p, g > =: \sum_{r_1+\ldots+r_m=n} a_{r_1,\ldots,r_m} b_{r_1,\ldots,r_m} \geq AB \frac{vdw(nm)}{vdw(n)^m} \quad (12)$$

Define

$$< p, g >_F =: \sum_{r_1+\ldots+r_m=n} a_{r_1,\ldots,r_m} b_{r_1,\ldots,r_m} (r_1)!\ldots(r_m)!$$

Is it true that (even for $l_i = \frac{n}{m}, 1 \leq i \leq m$ and multilinear polynomials)

$$< p, g >_F \geq AB \frac{n!}{m^n}.$$

The reason: the Van Der Waerden conjecture for the permanent sharply quantifies Hall's theorem on the rank of the intersection of two transveral matroids, the Bapat's conjecture sharply quantifies Rado's theorem on the rank of the intersection of one transversal and one geometric matroid.

The inequality (12) does similar thing for the intersection of two geometric matroids (even for the intersection of sets of vertices of special integer polymatroids). But the inequality (12), although non-obvious and quite cool, does not seem sharp.

A few "optimizational" comments.

1. Let $p \in Hom_+(n,n)$ be **H-Stable**

$$p(x_1, ..., x_n) = \sum_{r_1+...+r_n=n} a_{r_1,...,r_n} \prod_{1 \leq i \leq m} x_i^{r_i}.$$

   Existence and uniqueness in $\inf_{y_1+...+y_n=0} \log(p(e^{y_1}, ..., e^{y_n}))$:

$$a_{2,0,1,...,1}, a_{0,2,1,...,1}, ..., a_{1,...,1,2,0}, a_{1,...,1,0,2} > 0. \quad (13)$$

   I.e. $n(n-1)$ conditions, can be checked by a deterministic Strongly poly-time Black-Box algorithm; provides good bounds on a ball containing the solution.

2. If a minimum exists then either (13) hold or there exists a partition $\{1, ..., n\} = \cup_{1 \leq j \leq k} X_j$ such that

$$p(x_1, ..., x_n) = \prod_{1 \leq j \leq k} p_j(x_i, i \in X_j).$$

   This factorization can be also effectively computed.

3. Weak log-concavity: $p^{\frac{1}{n}}$ is concave on half-lines $\{X + tY : t \geq 0\} : X, Y \in R^n_{++}$.

Implies the inequality:

$p(SH(x_1,...,x_n)) \leq p(x_1,...,x_n)$, where $x_i > 0, 1 \leq i \leq n$ and

$$SH(x_1,...,x_n) = (y_1,...,y_n) : y_i = \frac{f(x_1,...,x_n)}{\frac{\partial}{\partial x_i} f(x_1,...,x_n)}.$$

Corollary 0.10: *Let* $p \in Hom_+(n,n)$ *be weakly log-concave. Suppose that* $Cap(f) > 0$*,*

$\log(Cap(f)) \leq \log(f(x_1,...,x_n)) \leq \log(Cap(f)) + \epsilon$*;* $\epsilon \leq \frac{1}{10}$

and $\Pi_{1 \leq i \leq n} x_i = 1$*.*

Then

$$\sum_{1 \leq i \leq n} (1 - \frac{x_i \frac{\partial}{\partial x_i} f(x_1,...,x_n)}{f(x_1,...,x_n)})^2 \leq 10\epsilon \ . \qquad (14)$$